

CW



PCT

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/069,883

DATE: 04/15/2003^{8.6}
TIME: 14:06:32

Input Set : A:\10069883.txt

Output Set: N:\CRF4\04152003\J069883.raw

3 <110> APPLICANT: MALLET , Francois
 4 COSSET , Francois-loic
 5 BLOND , Jean-Luc
 6 LAVILLETTE , Dimitri
 7 BOUTON , Olivier
 8 RUGGIERI , Alessia
 10 <120> TITLE OF INVENTION: Method for detecting the expression of an envelope protein
 of a human
 11 endogenous retrovirus and uses of a gene coding for said protein
 13 <130> FILE REFERENCE: 112062
 15 <140> CURRENT APPLICATION NUMBER: 10/069,883
 16 <141> CURRENT FILING DATE: 2002-05-01
 18 <150> PRIOR APPLICATION NUMBER: PCT/FR00/02429
 19 <151> PRIOR FILING DATE: 2000-09-01
 21 <150> PRIOR APPLICATION NUMBER: FR-9911141
 22 <151> PRIOR FILING DATE: 1999-09-01
 24 <150> PRIOR APPLICATION NUMBER: FR-9911793
 25 <151> PRIOR FILING DATE: 1999-09-15
 27 <160> NUMBER OF SEQ ID NOS: 64
 29 <170> SOFTWARE: PatentIn version 3.1
 31 <210> SEQ ID NO: 1
 32 <211> LENGTH: 538
 33 <212> TYPE: PRT
 34 <213> ORGANISM: Homo sapiens
 36 <400> SEQUENCE: 1
 38 Met Ala Leu Pro Tyr His Ile Phe Leu Phe Thr Val Leu Leu Pro Ser
 39 1 5 10 15
 42 Phe Thr Leu Thr Ala Pro Pro Pro Cys Arg Cys Met Thr Ser Ser Ser
 43 20 25 30
 46 Pro Tyr Gln Glu Phe Leu Trp Arg Met Gln Arg Pro Gly Asn Ile Asp
 47 35 40 45
 50 Ala Pro Ser Tyr Arg Ser Leu Ser Lys Gly Thr Pro Thr Phe Thr Ala
 51 50 55 60
 54 His Thr His Met Pro Arg Asn Cys Tyr His Ser Ala Thr Leu Cys Met
 55 65 70 75 80
 58 His Ala Asn Thr His Tyr Trp Thr Gly Lys Met Ile Asn Pro Ser Cys
 59 85 90 95
 62 Pro Gly Gly Leu Gly Val Thr Val Cys Trp Thr Tyr Phe Thr Gln Thr
 63 100 105 110
 66 Gly Met Ser Asp Gly Gly Gly Val Gln Asp Gln Ala Arg Glu Lys His
 67 115 120 125
 70 Val Lys Glu Val Ile Ser Gln Leu Thr Arg Val His Gly Thr Ser Ser
 71 130 135 140
 74 Pro Tyr Lys Gly Leu Asp Leu Ser Lys Leu His Glu Thr Leu Arg Thr

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/069,883

DATE: 04/15/2003

TIME: 14:06:32

Input Set : A:\10069883.txt

Output Set: N:\CRF4\04152003\J069883.raw

```

75 145          150          155          160
78 His Thr Arg Leu Val Ser Leu Phe Asn Thr Thr Leu Thr Gly Leu His
79          165          170          175
82 Glu Val Ser Ala Gln Asn Pro Thr Asn Cys Trp Ile Cys Leu Pro Leu
83          180          185          190
86 Asn Phe Arg Pro Tyr Val Ser Ile Pro Val Pro Glu Gln Trp Asn Asn
87          195          200          205
90 Phe Ser Thr Glu Ile Asn Thr Thr Ser Val Leu Val Gly Pro Leu Val
91          210          215          220
94 Ser Asn Leu Glu Ile Thr His Thr Ser Asn Leu Thr Cys Val Lys Phe
95 225          230          235          240
98 Ser Asn Thr Thr Tyr Thr Thr Asn Ser Gln Cys Ile Arg Trp Val Thr
99          245          250          255
102 Pro Pro Thr Gln Ile Val Cys Leu Pro Ser Gly Ile Phe Phe Val Cys
103          260          265          270
106 Gly Thr Ser Ala Tyr Arg Cys Leu Asn Gly Ser Ser Glu Ser Met Cys
107          275          280          285
110 Phe Leu Ser Phe Leu Val Pro Pro Met Thr Ile Tyr Thr Glu Gln Asp
111          290          295          300
114 Leu Tyr Ser Tyr Val Ile Ser Lys Pro Arg Asn Lys Arg Val Pro Ile
115 305          310          315          320
118 Leu Pro Phe Val Ile Gly Ala Gly Val Leu Gly Ala Leu Gly Thr Gly
119          325          330          335
122 Ile Gly Gly Ile Thr Thr Ser Thr Gln Phe Tyr Tyr Lys Leu Ser Gln
123          340          345          350
126 Glu Leu Asn Gly Asp Met Glu Arg Val Ala Asp Ser Leu Val Thr Leu
127          355          360          365
130 Gln Asp Gln Leu Asn Ser Leu Ala Ala Val Val Leu Gln Asn Arg Arg
131          370          375          380
134 Ala Leu Asp Leu Leu Thr Ala Glu Arg Gly Gly Thr Cys Leu Phe Leu
135 385          390          395          400
138 Gly Glu Glu Cys Cys Tyr Tyr Val Asn Gln Ser Gly Ile Val Thr Glu
139          405          410          415
142 Lys Val Lys Glu Ile Arg Asp Arg Ile Gln Arg Arg Ala Glu Glu Leu
143          420          425          430
146 Arg Asn Thr Gly Pro Trp Gly Leu Leu Ser Gln Trp Met Pro Trp Ile
147          435          440          445
150 Leu Pro Phe Leu Gly Pro Leu Ala Ala Ile Ile Leu Leu Leu Leu Phe
151          450          455          460
154 Gly Pro Cys Ile Phe Asn Leu Leu Val Asn Phe Val Ser Ser Arg Ile
155 465          470          475          480
158 Glu Ala Val Lys Leu Gln Met Glu Pro Lys Met Gln Ser Lys Thr Lys
159          485          490          495
162 Ile Tyr Arg Arg Pro Leu Asp Arg Pro Ala Ser Pro Arg Ser Asp Val
163          500          505          510
166 Asn Asp Ile Lys Gly Thr Pro Pro Glu Glu Ile Ser Ala Ala Gln Pro
167          515          520          525
170 Leu Leu Arg Pro Asn Ser Ala Gly Ser Ser
171          530          535

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/069,883

DATE: 04/15/2003

TIME: 14:06:32

Input Set : A:\10069883.txt

Output Set: N:\CRF4\04152003\J069883.raw

```

174 <210> SEQ ID NO: 2
175 <211> LENGTH: 2781
176 <212> TYPE: DNA
177 <213> ORGANISM: Homo sapiens
179 <400> SEQUENCE: 2
180 atgggagctg ttttcatgct atttcaactct attaaatctt gcaactgcac tcttctggtc      60
182 catgtttctt acggctcgag ctgagctttt gctcaccgtc caccactgct gtttgccacc      120
184 accgcagacc tgccgctgac tcccatccct ctggatcctg cagggtgtcc gctgtgctcc      180
186 tgatccagcg aggcgcccac tgccgctccc aattgggcta aaggcttgcc attgttcctg      240
188 cacggctaag tgccctgggtt tgttctaatt gagctgaaca ctagtcaactg ggttccatgg      300
190 ttctcttctg tgacccacgg cttctaatag aactataaca cttaccacat ggcccaagat      360
192 tccattccctt ggaatccgtg aggccaaaga ctccagggtca gagaatacga ggcttgccac      420
194 catcttgga ggcgcctgct accatcttgg aagtgggttca ccaccatctt gggagctctg      480
196 tgagcaagga ccccccggtt acattttggc aaccacgaac ggacatccaa agtgatacat      540
198 cctgggaagg accctaccca gtcattttat ctaccccaac tgcggttaaa gtggctggag      600
200 tggagctctg gatacatcac acttgagtca aatcctggat actgccaaag gaacctgaaa      660
202 atccaggaga caacgctagc tattcctgtg aacctctaga ggatttgccg ctgctcttca      720
204 aacaacaacc aggaggaaa gtaactaaaat cataaatccc catggccctc ccttatcata      780
206 tttttctctt tactgttctt ttacctctt tcaactctac tgcacccctc ccatgccgct      840
208 gtatgaccag tagtcccctt taccaagagt ttctatggag aatgcagcgt cccggaaata      900
210 ttgatgcccc atcgatatagg agtctttcta agggaacccc caccttcaact gccacacccc      960
212 atatgccccg caactgctat cactctgcca ctctttgcat gcatgcaaat actcattatt      1020
214 ggacaggaaa aatgattaat cctagtgtgc ctggaggact tggagtcaact gtctgttgga      1080
216 cttacttcac caaaactggt atgtctgatg ggggtggagt tcaagatcag gcaagagaaa      1140
218 aacatgtaaa agaagtaatc tcccaactca cccgggtaca tggcacctct agccctaca      1200
220 aaggactaga tctctcaaaa ctacatgaaa ccctccgtac ccatactcgc ctggtaagcc      1260
222 tatttaatac caccctcact gggctccatg aggtctcggc ccaaaaccct actaactgtt      1320
224 ggatatgcct cccctgaac ttcaggccat atgtttcaat ccctgtacct gaacaatgga      1380
226 acaacttcag cacagaaata aacaccactt ccgttttagt aggacctctt gtttccaatc      1440
228 tggaaataac ccatacctca aacctcacct gtgtaaaatt tagcaatact acatacaca      1500
230 ccaactccca atgcatcagg tgggtaactc ctcccacaca aatagtctgc ctacctcag      1560
232 gaatatTTTT tgtctgtggt acctcagcct atcgttgttt gaatggctct tcagaatcta      1620
234 tgtgcttctc ctacttctta gtgcccccta tgaccatcta cactgaacaa gatttatata      1680
236 gttatgtcat atctaagccc cgcaacaaaa gagtaccat tcttctttt gttataggag      1740
238 cgggagtgtc aggtgcaacta ggtactggca ttggcggtat cacaacctct actcagttct      1800
240 actacaaact atctcaagaa ctaaatgggg acatggaacg ggtcgccgac tccttggtca      1860
242 ccttgcaaga tcagcttaac tccctagcag cagtgtcct tcaaaatcga agagctttag      1920
244 acttgctaac cgctgaaaga gggggaacct gtttattttt aggggaagaa tgctgttatt      1980
246 atgttaatca atccggaatc gtcactgaga aagttaaaga aattcgagat cgaatacaac      2040
248 gtagagcaga ggagcttcga aacactggac cctggggcct cctcagccaa tggatgcct      2100
250 ggattctccc cttcttagga cctctagcag ctataatatt gctactctc tttggaccct      2160
252 gtatctttaa cctccttggt aactttgtct cttccagaat cgaagctgta aaactacaaa      2220
254 tggagcccaa gatgcagtcc aagactaaga tctaccgag acccctggac cggcctgcta      2280
256 gccacgatac tgatgttaat gacatcaaag gcacccctcc tgaggaaatc tcagctgcac      2340
258 aacctctact acgccccaat tcagcaggaa gcagttagag cggctcgtcg ccaacctccc      2400
260 caacagcact taggttttcc tgttgagatg ggggactgag agacaggact agctggattt      2460
262 cctaggctga ctaagaatcc ctaagcctag ctgggaagg gaccacatcc acctttaaac      2520
264 acggggcttg caacttagct cacacctgac caatcagaga gctcactaaa atgctaatta      2580
266 ggcaaagaca ggaggtaaag aaatagccaa tcatctattg octgagagca cagcaggagg      2640

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/069,883

DATE: 04/15/2003

TIME: 14:06:32

Input Set : A:\10069883.txt

Output Set: N:\CRF4\04152003\J069883.raw

```

268 gacaatgatac gggatataaa cccaagtctt cgagccggca acggcaaccc cctttgggtc 2700
270 ccctcccttt gtatgggagc tctgttttca tgctatttca ctctattaaa tcttgcaact 2760
272 gcaaaaaaaaa aaaaaaaaaa a 2781
275 <210> SEQ ID NO: 3
276 <211> LENGTH: 28
277 <212> TYPE: DNA
278 <213> ORGANISM: Homo sapiens
280 <400> SEQUENCE: 3
281 caaaacgcct ggagatacag caattatc 28
284 <210> SEQ ID NO: 4
285 <211> LENGTH: 25
286 <212> TYPE: DNA
287 <213> ORGANISM: Homo sapiens
289 <400> SEQUENCE: 4
290 gcaccctcat ggttgtgtta cttgg 25
293 <210> SEQ ID NO: 5
294 <211> LENGTH: 26
295 <212> TYPE: DNA
296 <213> ORGANISM: Homo sapiens
298 <400> SEQUENCE: 5
299 ctgaaaatcc aggagacaac gctagc 26
302 <210> SEQ ID NO: 6
303 <211> LENGTH: 25
304 <212> TYPE: DNA
305 <213> ORGANISM: Homo sapiens
307 <400> SEQUENCE: 6
308 gcaccctcat ggttgtgtta cttgg 25
311 <210> SEQ ID NO: 7
312 <211> LENGTH: 36
313 <212> TYPE: DNA
314 <213> ORGANISM: Homo sapiens
316 <400> SEQUENCE: 7
317 ttggtaccca aaacgcctgg agatacagca attatc 36
320 <210> SEQ ID NO: 8
321 <211> LENGTH: 29
322 <212> TYPE: DNA
323 <213> ORGANISM: Homo sapiens
325 <400> SEQUENCE: 8
326 aactcgagtg aaatagcatg aaaacagag 29
329 <210> SEQ ID NO: 9
330 <211> LENGTH: 25
331 <212> TYPE: DNA
332 <213> ORGANISM: Homo sapiens
334 <400> SEQUENCE: 9
335 aggaaagtaa ctaaaatcat aaatc 25
338 <210> SEQ ID NO: 10
339 <211> LENGTH: 20
340 <212> TYPE: DNA
341 <213> ORGANISM: Homo sapiens

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/069,883

DATE: 04/15/2003

TIME: 14:06:32

Input Set : A:\10069883.txt

Output Set: N:\CRF4\04152003\J069883.raw

```

343 <400> SEQUENCE: 10
344 ggttccctta gaaagactcc 20
347 <210> SEQ ID NO: 11
348 <211> LENGTH: 22
349 <212> TYPE: DNA
350 <213> ORGANISM: Homo sapiens
352 <400> SEQUENCE: 11
353 aatattgatg ccccatcgta ta 22
356 <210> SEQ ID NO: 12
357 <211> LENGTH: 21
358 <212> TYPE: DNA
359 <213> ORGANISM: Homo sapiens
361 <400> SEQUENCE: 12
362 ccagtttggg tgaagtaagt c 21
365 <210> SEQ ID NO: 13
366 <211> LENGTH: 21
367 <212> TYPE: DNA
368 <213> ORGANISM: Homo sapiens
370 <400> SEQUENCE: 13
371 ggaggacttg gagtcactgt c 21
374 <210> SEQ ID NO: 14
375 <211> LENGTH: 20
376 <212> TYPE: DNA
377 <213> ORGANISM: Homo sapiens
379 <400> SEQUENCE: 14
380 aggcgagtat gggtagcgag 20
383 <210> SEQ ID NO: 15
384 <211> LENGTH: 23
385 <212> TYPE: DNA
386 <213> ORGANISM: Homo sapiens
388 <400> SEQUENCE: 15
389 ggactagatc tctcaaaact aca 23
392 <210> SEQ ID NO: 16
393 <211> LENGTH: 22
394 <212> TYPE: DNA
395 <213> ORGANISM: Homo sapiens
397 <400> SEQUENCE: 16
398 acggaagtgg tgtttatttc tg 22
401 <210> SEQ ID NO: 17
402 <211> LENGTH: 21
403 <212> TYPE: DNA
404 <213> ORGANISM: Homo sapiens
406 <400> SEQUENCE: 17
407 cctgaacaat ggaacaactt c 21
410 <210> SEQ ID NO: 18
411 <211> LENGTH: 20
412 <212> TYPE: DNA
413 <213> ORGANISM: Homo sapiens
415 <400> SEQUENCE: 18

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/069,883

DATE: 04/15/2003
TIME: 14:06:33

Input Set : A:\10069883.txt
Output Set: N:\CRF4\04152003\J069883.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:54; Xaa Pos. 129,224,307
Seq#:55; Xaa Pos. 129,224,307

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/069,883

DATE: 04/15/2003

TIME: 14:06:33

Input Set : A:\10069883.txt

Output Set: N:\CRF4\04152003\J069883.raw

L:820 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:54,Line#:817
L:854 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:54 after pos.:432
M:341 Repeated in SeqNo=54
L:1012 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55 after pos.:128
M:341 Repeated in SeqNo=55
L:1127 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:56,Line#:1124
L:1419 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:58,Line#:1416
L:1711 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:60,Line#:1708